Explaining survey nonresponse in the ESS in Belgium using municipality-level administrative registry data

Celine Wuyts,* Centre for Sociological Research, KU Leuven

Anina Vercruyssen, Centre for Sociological Research, KU Leuven

Geert Loosveldt, Centre for Sociological Research, KU Leuven

* Corresponding author: Celine Wuyts, Centre for Sociological Research, KU Leuven, Leuven, Belgium. Tel.: +32 16 37 70 30; Fax: +32 16 3 23365; E-mail: celine.wuyts@soc.kuleuven.be.

Abstract

Suitable auxiliary data is important in order to assess the potentially detrimental effects of nonresponse on survey estimates. Whereas sufficient individual-level auxiliary data is rarely available, aggregated data is often quite readily accessible. We investigate whether municipality-level data from official administrative registries is useful to explain and predict individuals' survey response outcomes in two rounds of the European Social Survey (ESS) in Belgium. This study was prompted by the recent publication of the 2011 Belgian Census. This is the first Belgian census produced by linking various national administrative databases instead of a nationwide survey. It offers free and easy access to a broad range of data at the municipality level. We find no consistent results in the two rounds of the ESS, and the overall usefulness of aggregated data appears limited. Individual-level data available from other auxiliary sources is more strongly and more consistently related to survey response outcomes. Nonresponse analyses in the ESS and other surveys would benefit from access to individual-level administrative registry data, but as this is usually restricted, alternative auxiliary data sources such as interviewer observations will need to suffice.

Keywords: nonresponse, auxiliary data, aggregated, regional characteristics, European Social Survey

1 Introduction

Observed declines in survey response rates [1] and challenges experienced by surveys such as the European Social Survey in their pursuit of ambitious response rate targets [2] [3] have encouraged research into the determinants of nonresponse, with the aims of understanding the underlying mechanisms, adapting survey design features and fieldwork decisions to limit nonresponse error, and/or statistically adjusting for nonresponse bias. Whatever the goal, suitable auxiliary variables are required that are available for all sample units: Respondents and non-respondents. Finding appropriate auxiliary variables can be challenging, because they should relate to the probability of survey participation and ideally to survey variables of interest [4], and different sources of auxiliary variables are associated with different costs and quality [5].

Potential auxiliary variables are occasionally directly available from the sampling frame. Much more commonly, however, they have to be retrieved from external sources. Individual or household-level auxiliary data from external sources has contributed substantially to a better understanding of nonresponse mechanisms (see e.g., Groves & Couper, 1998 [6] and Durrant & Steele, 2009 [7]). Both contactability and reluctance to cooperate are known to be associated with observable characteristics of sample units, such as family composition and employment status, even if these observable characteristics are only rough proxies of people's socio-psychological dispositions [8]. However, procuring such low-level linkable data is somewhat demanding, costly, and hindered by privacy regulations. Data aggregated to a regional level, which can be assigned to both respondents and non-respondents, is often more readily available to survey researchers and practitioners, and has also been shown to relate to survey response outcomes.

One regional correlate of survey nonresponse that is both intuitive and supported by consistent empirical evidence is urbanicity. People in more urbanized areas have been found to be less likely to participate in surveys (see e.g., Groves & Couper, 1998 [6] and House & Wolf, 1978 [9]), which may be due to crowding and stimulus overload, perceived risk of crime, and/or (lack of) social cohesion [6]. The social cohesion theory hinges on the assumption that survey participation is a form of social involvement [10] [11], and that weakened social ties and institutions limit people's willingness to cooperate. However, empirical support for the social cohesion theory is ambiguous. Groves and Couper (1998) [6], for example, found a positive association between survey cooperation and the proportion of young people in a community, whereas no such association was observed for other potential indicators of community-level social cohesion. Johnson and colleagues (2006) [12] found that refusal to participate in a telephone survey in the state of Illinois was more likely in areas characterized by concentrated affluence and in residentially stable areas, even though residential stability is thought to facilitate rather than hamper the formation of social ties [13].

In line with earlier research mentioned in the previous paragraph, in which regional characteristics are linked to individual survey outcomes, we investigate the potential relevance of municipality-level data that are easily available from official administrative registries with regard to explaining and predicting individuals' survey response outcomes in the European Social Survey (ESS) in Belgium.

A necessary first step in this type of study consists of identifying relevant components of regional differentiation: In what way do neighborhoods or larger geographic areas differ so that survey nonresponse is more likely or less likely? Most previous research (see e.g., House & Wolf, 1978 [9]; Groves & Couper, 1998 [6]; Johnson et al., 2006 [12]; Matsuoka & Maeda, 2015 [14]) takes a theory-driven approach to this step and our first research question builds on this method. We turn to research on the collective efficacy of the social environment, which captures both social cohesion and willingness to contribute to the common good [13], and which is hypothesized to affect people's survey participation decisions [12]. How do the structural components of regional differentiation related to collective efficacy, as identified by Sampson and colleagues (1999) [13] – residential stability, concentrated affluence, concentrated disadvantage, urbanicity, concentrated immigration, and emphasis on children and family – relate to survey response outcomes in the Belgian ESS? In order to answer this question, we essentially attempt to replicate the analysis of regional correlates of survey nonresponse carried out by Johnson and colleagues (2006) [12].

An alternative approach to identify potentially relevant regional differentiators is data driven, by first applying a dimension reduction procedure, such as exploratory factor analysis, to discover the underlying dimensions that capture most of the regional variation (e.g. Van Goor, Jansma, & Veenstra, 2005 [15]). This is the approach we take in order to find an answer to our second research question: Which are the main dimensions underlying the observed municipality-level variation in the official registry data, and how do they relate to survey response outcomes in the Belgian ESS?

2 Data and methods

To examine the potential relevance of municipality-level data from official administrative registries with regard to nonresponse analysis, we use individual-level response outcome data from rounds 6 and 7 of the European Social Survey (ESS) in Belgium. The ESS is a biennial, cross-sectional survey on social topics such as migration, democracy, welfare, and well-being. The ESS aims for a 70% response rate in all countries but many countries, including Belgium, have continuously fallen short of the target. Even so, response rates have remained relatively stable in the majority of other ESS countries including Belgium [3]. Because the ESS aims to facilitate not only cross-national but also intertemporal comparisons, its key design features are stable over different rounds. In particular, the mode of data collection (face-to-face interviews) and the contact procedure (at least four personal visits spread over different days of the week, different times of day, and over at least two different weeks). Further, a large part of the ESS questionnaire is identical in the two rounds considered [16] [17].

In the case of Belgium, the agency employed to implement the fieldwork is the same for both rounds, and consequently the interviewer workforce in the two rounds strongly overlaps (88 out of the 150 interviewers active in ESS round 7 had been actively involved in round 6) and all interviewers were briefed on the project in a similar fashion. For round 6 (fieldwork in the fall of 2012), 3,267 individuals were randomly selected from the Belgian National Registry using a two-stage clustered sampling design. In the first stage, 221 of the 589 Belgian municipalities were randomly selected. In the second stage, individuals were drawn from the selected municipalities with the number in each municipality proportional to its population size. For round 7 (fieldwork in the fall of 2014), 3,204 individuals were drawn in the same manner, again from 221 municipalities.

Although the intended implementation is comparable, the fieldwork process differed in some profound ways. The ESS round 6 fieldwork got off to a flying start and overall progressed relatively smoothly. The fieldwork for round 7, on the other hand, commenced with some difficulty and slowed down substantially after a month and a half, in the end necessitating an extension of the fieldwork period and more extensive refusal conversion efforts to achieve a roughly similar final response rate.

We analyze the round 6 and round 7 samples separately, but in an identical way as replicate samples in order to evaluate the robustness of the results. Even though the round 7 fieldwork process may have been less homogeneous, and therefore more strongly affected by regional differences in composition and social environment, the municipal-level variables are expected to relate to the individual-level response outcomes in a similar way in the two rounds.

Outcome variables

Individual-level survey response outcome data from the ESS round 6 and 7 [18] [19] in Belgium was categorized into respondents, noncontacts, refusals, not able and other nonresponse, and ineligible. Based on this categorization, we constructed three binary indicators for the final response outcome: (a) whether an eligible sample unit is a noncontact case (noncontact), (b) whether a contacted sample unit is a refusal case (refusal, given contact), and for overall nonresponse irrespective of the source, (c) whether an eligible sample unit is a completed interview case. The final response outcome distributions and outcome rates are presented in Table 1.

Municipality-level independent variables

A large number of municipality-level variables are available from Statistics Belgium. The majority of variables originate from the 2011 Belgian Census [20], for the first time constructed by linking various administrative registries and completed at the end of 2014, complemented with annual statistics. This decennial census provides information on the composition of municipalities with regard to residents' age, gender, nationality, marital status, family structure, occupation, and the type and age of housing. Total population counts, derived from the Belgian National Registry, and fiscal statistics such as median income and income inequality, derived from personal income tax declarations, are published yearly; We use the 2012 data. This data represents the Belgian population in 2011/2012, but can be assumed to be relatively stable and can therefore be reasonably linked to the individual-level response outcomes of the ESS round 6 (2012) and 7 (2014).

Given the number of municipality-level variables available and the strong correlations that can be expected between them [9] [14], some type of dimension reduction approach is required. One way to reduce the huge number of variables in the data is a theory-driven selection, based on existing literature. We operationalize the six components of regional differentiation identified by Sampson and colleagues (1999) [13], and in a manner similar to that used by Johnson and colleagues (2006) [12]. Specifically: (1) the percentage of address changes as an indicator of residential (in)stability, (2) the percentage of employers in the working population as an indicator of concentrated affluence, (3) whether or not the median yearly income exceeds EUR 35,000 (approximately the first quartile of the median income distribution) as an indicator of concentrated disadvantage, (4) the population density as an indicator of urbanicity, (5) the percentage of people with non-Belgian nationality as an indicator of concentrated immigration, and (6) the percentage of children younger than 15 as an indicator of emphasis on children and family. Table 2 shows the Pearson correlation matrix for these six indicators.

Another way to reduce the number of variables in the municipality-level data is to extract a limited set of (uncorrelated) dimensions using factor analysis. The main idea of factor analysis is that there are some underlying dimensions (factors), which produce the observed correlations between variables. The aim of exploratory factor analysis is to reveal these underlying dimensions, and to represent the data in a smaller number of variables whilst limiting the loss of information. In the first step, we selected 49 variables based on preliminary descriptive analysis. Variables with little variability were either dropped (e.g. 'the proportion of people with Oceanic nationality in the non-Belgium population') or combined with others (e.g. 'the proportion of widows(ers) previously in registered partnerships' combined with 'the proportion of widows(ers) previously married'). The remaining municipality-level variables, along with some basic descriptive statistics, are presented in Table 3. Together, these variables describe the structure of the Belgian municipalities in terms of population density, income (inequality), socio-demographics, and housing. Many of them are significantly correlated. Some observed correlations confirm expected associations, such as population density and the proportion of houses in multi-unit buildings ($r = .62$, $p < .0001$). Other observed correlations do not have an obvious explanation, for example the proportion of people who changed their address and the proportion of homemakers ($r = .67$, $p < .0001$). Somehow, these correlations suggest the existence of underlying dimensions causing the observed correlational pattern. We subsequently summarized the 49 municipality-level variables in a more practical number of uncorrelated municipality dimensions by using maximum-likelihood exploratory factor analysis. The Kaiser-Meyer-Olkin measure of factor adequacy was .71, which is sufficiently high for factor analysis to be deemed suitable. The result of the factor analysis is a simpler structure that still captures a large proportion of the variability in the municipality-level data.

Because only a subset of municipalities are selected in each ESS round and this subset only partially overlaps between the two rounds (46.6%), we verified that the subset of municipalities selected in the two rounds is not significantly differently distributed with regard to the six single indicators and the dimensions derived from factor analysis.

Whereas a theory-driven selection approach restricts attention to a limited set of variables, exploratory factor analysis benefits from covering the available data more generally. The other side of the coin is that variables selected based on theory are specific and well defined, whereas exploratory factor analysis tends to produce a set of explanatory variables with more ambiguous interpretations. We also recognize that there is some subjectivity in both approaches, as potential candidate indicators of theoretical constructs may vary with regard to validity, accuracy, and predictive power, and extracted factors usually do not correspond well with theoretical constructs. Further, the factor structure can be interpreted in different ways.

Individual-level independent variables

Two individual-level variables are directly available from the sampling frame: The age and gender of the sample unit. We use six categories for age (24 years or younger, 25 to 34 years, 35 to 44 years, 45 to 54 years, 55 to 64 years, and 65 years or older) to allow for unequal age effects on survey response outcomes. Therefore, five age category dummies are included, with the group of sample units aged 24 or younger used as the reference group. Interviewers' observations concerning the housing situation of the sample units (normally collected in the ESS) add two additional individual-level variables, specifically the presence of impediments to access and the physical condition of housing. The first is included as a binary variable indicating whether an entry phone and/or locked gate impede direct access to the sample unit's front door. The latter is an ordinal variable included as numeric, ranging from 1 ('Very good') to 5 ('Very bad'). Although the sampling frame variables, age and gender, are available for all sample units, interviewer observations are missing for 29 sampling units in round 6 and nine in round 7. The number of complete observations (excluding ineligible sample units) available for analysis is 3,160 for round 6 and 3,095 for round 7 (Table 1).

Modeling approach

For each of the three binary dependent variables – overall survey nonresponse, noncontact, and refusal conditional on contact – we estimate a logistic multilevel model with sample units nested in municipalities. The sequential modelling approach of noncontact and refusal, given contact, is the most common approach in the nonresponse modelling literature [21]. This approach recognizes the two-stagedness of the survey response process [6] [22] and the parameter estimates are intuitively interpretable.

An intercept-only model (Model 0), containing no explanatory variables, is used to estimate the intraclass correlation coefficient (ICC), indicating the proportion of variability in the dependent variable attributable to differences between municipalities. As the outcome variable is binary, the residual variance component is fixed at $\frac{\pi^2}{3}$ [23]. In the next step, we add two different sets of municipality-level variables. Specifically, the set of six single indicators (percentage of address changes, percentage of employers in the working population, whether or not the median yearly income exceeds EUR 35,000, population density, percentage of people with a non-Belgian nationality, and percentage of children younger than 15) for the theory-driven selection approach, and the set of factors derived from exploratory factor analyses for the data-driven approach. In the final step, we add the individual-level variables, including age and gender obtained from the sampling frame, and interviewer observations on the presence of impediments to access and the condition of housing.

3 Results

The results of the two approaches to modelling survey nonresponse as a function of municipality-level administrative registry data, namely theory-based selection of single indicators and dimension extraction using factor analysis, for the ESS round 6 and 7 in Belgium, are presented in the following subsections.

Before turning to the main results we evaluate the extent of municipality-level differences in survey outcome rates. The intraclass correlation coefficients estimated by the intercept-only models suggest some systematic differences in outcome rates between municipalities, especially for noncontact rates. Between 10.0% (round 7) and 13.7% (round 6) of the variance in noncontact is estimated to be at the municipality level, whereas only between 4.9% (round 7) and 5.9% (round 6) of the variance in refusal, given contact, is at this level. These estimates further suggest that municipality-level differences are somewhat less important in round 7 than in round 6 of the ESS.

3.1 Nonresponse analysis with municipality variables selected based on theory

In the first analysis, we examine the association between municipality-level single indicators of regional differentiation derived from theory (percentage of address changes, percentage of employers in the working population, whether or not the median yearly income exceeds EUR 35,000, population density, percentage of people with a non-Belgian nationality, and percentage of children younger than 15) and individuals' response outcomes in the ESS for Belgium. The coefficient estimates of the municipality-level indicators on the probability of survey nonresponse, noncontact, and refusal, given contact, in round 6 of the ESS for Belgium are shown in Table 4. The estimates for round 7 are shown in Table 5.

The results suggest that some of the components of regional differentiation are related to survey response outcomes in the ESS, but no consistent pattern can be observed over the two rounds. In round 6, overall survey nonresponse does not relate to any of the municipality-level indicators (Table 4, Model 1). Taking a closer look at the two main components of nonresponse – nonresponse due to noncontact and nonresponse due to refusal – we find no statistically significant effect of any of the municipality-level indicators on the probability of noncontact, and only one on the probability of refusal, given contact. Refusal is significantly less likely in low-income municipalities. By contrast, noncontact is somewhat more likely in low-income municipalities. The coefficient estimate is not significantly different from zero at the 5% significance level, but the effect is large enough to offset the positive effect of living in a low-income municipality on overall survey response. In order to test the relevance of the municipality-level correlates of nonresponse over and above the available individual-level data, in the final step of the analysis we add individuals' age, gender, impediments to access, and housing conditions. The municipality-level indicators' coefficient estimates change only slightly (Table 4, Model 2). Refusal remains significantly less likely in low-income municipalities, even after adding the individuals' basic socio-demographics and interviewer-observed housing characteristics.

The results for round 7 (Table 5, Model 1) are strongly different. We observe no effect of living in a low-income municipality on survey response outcomes. Instead, two other municipality-level indicators appear to relate to overall survey nonresponse: Population density and the percentage of children. People living in municipalities with a higher population density and fewer children are significantly more likely to be non-respondents. Again, we distinguish between nonresponse due to noncontact and nonresponse due to refusal, and find no statistically significant effect of any of the municipality-level indicators on the probability of noncontact. Refusal, on the other hand, is significantly more likely in municipalities with fewer children, fewer people with non-Belgian nationality, and a higher population density. Adding the individuals' age, gender, the presence of impediments to access, and housing conditions does not strongly alter the coefficient estimates of the municipality-level indicators (Table 5, Model 2). The effect of population density on the probability of refusal, however, ceases to reach the traditional 5% level of significance when we add the individual-level variables.

In Table 6, we present some goodness-of-fit criteria for each modeling step (Model 0, Model 1, Model 2) for each of the three outcome variables in the two rounds. In round 6, Model 0 (without any explanatory variables) is preferred over Model 1 (including the municipality-level single indicators of regional differentiation) for all three survey response outcome, both by the AIC and the more conservative BIC. The likelihood ratio test also suggests that no improvement in goodness-of-fit for either survey response outcome is achieved by including the indicators of regional differentiation. Including the individual-level age, gender and interviewer observations (Model 2), on the other hand, does significantly improve the goodness-of-fit. Model 0 remains nonetheless preferred by the BIC. In round 7, Model 1 is preferred over Model 0 for all three survey response outcome by the AIC but not by the BIC. The likelihood ratio test also suggests a significant improvement in goodness-of-fit for all three survey response outcomes from including the indicators of regional differentiation. Including the individual-level age, gender and interviewer observations (Model 2) further improves the goodness of fit. Again, Model 0 remains preferred by the BIC. These criteria suggest that only the round 7 response outcome models benefit from including the municipality-level single indicators.

3.2 Nonresponse analysis with municipality dimensions based on exploratory factor analysis

In the second analysis, we explore the dimensions underlying the observed municipality-level variation in the official registry data, and examine their association with individuals' survey response outcomes in the ESS for Belgium. Exploratory factor analysis on the municipality-level variables yields four uncorrelated factors. Together, they explain 59.5% of the variation in the available municipality data. The four-factor solution was preferred on the basis of the scree plot which suggested that eigen values started to level off quickly after four factors. The fifth and sixth factor explained only just over 5% of the variance. We use a varimax rotation to simplify the interpretation of the factors. We also considered promax rotations but the solutions are highly similar and we therefore prefer the four-factor varimax rotated solution for its intuitive appeal with uncorrelated factors. The factor loadings are shown in Table 3.

The first dimension explains 27.1% of the variation and is labeled ‘residential instability’ due to the high loadings by the following variables: The proportion of people in rented homes (.92), the proportion of homes in multi-unit buildings (.86), the proportion of single-person households (.84), and the proportion of people who changed their address (.81). The municipalities in the Brussels-Capital Region score especially high on this dimension. Other larger cities in Belgium also have positive scores, but not nearly as high as those in the Brussels area. The second dimension (labeled ‘concentration of elderly people) explains an additional 14.2% of the municipality variation. Variables loading strongly on this dimension are the proportion of people aged 65 or over (.93), the proportion of retired people (.92), and the proportion of couples without children (.85). The municipalities in the coastal area score high on this dimension, and municipalities in the Walloon Region tend to score somewhat lower than those in the Flemish Region. The third dimension (labeled ‘concentration of previously partnered people) explains an additional 11.1% of the municipality variation. This dimension is characterized by strong loadings of the proportion of widowed people (.75), divorced people (.64), and single parents with older children (.63). Remarkably, the proportion of older buildings also loads strongly on this dimension (.68). The fourth dimension (labeled ‘socioeconomic status’) explains an additional 7.3% of the municipality variation. The proportion of people with a degree of higher education and the median income load strongly on this dimension (.95 and .72, respectively). The scores are notably high in some of the municipalities that adjoin university cities. The municipalities in Flemish Brabant and Walloon Brabant in general tend to score above average, whereas the Brussels-Capital Region is divided, with low-scoring municipalities in the northwest and high-scoring municipalities in the southeast of the area.

The coefficient estimates for the four municipality dimensions on the probability of survey nonresponse, noncontact, and refusal in round 6 of the ESS are shown in Table 7. The estimates for round 7 are presented in Table 8. One municipality dimension is found to be negatively associated with survey response in both rounds, namely residential instability. There is also partial evidence (in round 7) that survey response is negatively affected by a high concentration of the elderly and the high socioeconomic status of neighborhoods. Because of the observed differences between round 6 and round 7, we describe the results separately.

In round 6 of the ESS (Table 7, Model 1), the probability of overall survey nonresponse is significantly higher in municipalities that score high on the 'residential instability' dimension, but does not relate to any of the other municipality dimensions. Taking a closer look at nonresponse due to noncontact and nonresponse due to refusal, we observe that noncontact is somewhat more likely in municipalities that are residentially unstable, but the coefficient estimate is not significantly different from zero at the 5% significance level. In addition, noncontact is less likely in municipalities with a higher concentration of elderly people. Even though this effect is statistically significant, it is not large enough to influence the probability of overall survey nonresponse: Overall survey nonresponse is not significantly more or less likely in municipalities with many elderly people. We do not find a statistically significant effect of any of the municipality dimensions on the probability of refusal. Again, in order to check the relevance of the municipality-level correlates of nonresponse over and above the available individual-level data, in the final step of the analysis we add individuals' age, gender, impediments to access, and housing conditions. Some of the coefficient estimates for the municipality dimensions change considerably (Table 7, Model 2): The effect of residential instability on overall survey nonresponse and the effect of the proportion of elderly people on noncontact cease to reach statistical significance, to the extent that none of the municipality dimensions are significantly associated with survey response outcomes in the ESS round 6 after adding the individual-level basic socio-demographics and interviewer-observed housing characteristics.

The results for round 7 of the ESS (Table 8, Model 1) show a more elaborate set of effects of municipality dimensions on survey response outcomes. The probability of overall survey nonresponse is significantly higher in municipalities that score high on the ‘concentration of elderly people’ and ‘socioeconomic status’ dimensions, as well as the ‘residential instability’ dimension. The results for nonresponse due to noncontact and nonresponse due to refusal indicate that residential instability has a negative effect on survey response for both. People living in more residentially unstable municipalities are both significantly less likely to be successfully contacted and significantly more likely to refuse participation in the survey. The results further suggest that socioeconomic status has a negative effect on survey response both through noncontact and refusal, and that the proportion of elderly people has a negative effect primarily through refusal. People living in municipalities with a higher socioeconomic status are significantly harder to reach, and significantly more likely to refuse participation even if successfully contacted. People living in municipalities with many elderly, on the other hand, are not significantly harder to reach, but are significantly more likely to refuse.

Similarly to round 6 of the ESS, some of the municipality dimensions’ coefficient estimates change very little and others change considerably when we add individuals’ age, gender, impediments to access, and housing conditions in the final step (Table 8, Model 2). The probability of overall survey nonresponse remains significantly higher in municipalities that score high for the ‘residential instability’, ‘concentration of elderly people’, and ‘socioeconomic status’ dimensions. Although the effect of residential instability on overall survey nonresponse remains significant and in the expected direction, its effect on the two components of survey nonresponse in round 7 cease to reach statistical significance at the 5% level. The findings that people living in municipalities with a high socioeconomic status are significantly less likely to be successfully contacted and significantly more likely to refuse, whereas people living in municipalities with many elderly are significantly more likely to refuse participation remain robust to additionally adding the individual-level variables.

Again we present some goodness-of-fit criteria for each modeling step (Model 0, Model 1, Model 2) for each of the three outcome variables in both rounds, now with the municipalities' dimensions instead of the municipality-level single indicators added in Model 1 (Table 9). In round 6, Model 1 (including the municipality dimensions) is (marginally) preferred over Model 0 (without any explanatory variables) for noncontact and overall nonresponse (but not for refusal, given contact) by the AIC, but not by the BIC. The likelihood ratio test suggests a significant improvement in goodness-of-fit from including the municipality dimensions only for noncontact. Including the individual-level age, gender and interviewer observations (Model 2), on the other hand, does significantly improve the goodness-of-fit for all three survey response outcomes. Model 0 remains nonetheless preferred by the BIC. In round 7, Model 1 (including the municipality dimensions) is preferred over Model 0 (without any explanatory variables) for all three survey response outcomes by the AIC, and for noncontact and (marginally) for refusal, given contact (but not for overall nonresponse) by the BIC. The likelihood ratio test also suggests a significant improvement in goodness-of-fit for all three survey response outcomes from including the municipality dimensions. Model 0 remains preferred by the BIC for refusal, given contact, but Model 2 is preferred for noncontact and overall nonresponse. These criteria suggest that the response outcome models benefit from including the municipality dimensions, but more so in round 7 than in round 6, and more so for nonresponse due to noncontact than for nonresponse due to refusal.

4 Conclusions and discussion

In this study, we explore the extent to which readily obtainable municipality-level data from official administrative registries, such as the 2011 Belgian Census, can be used to explain survey nonresponse and its two main components, noncontact and refusal, in the ESS. We illustrate two approaches to identifying components of regional differentiation, one driven by theoretical considerations and one driven by the variability observed in the data. For the first approach, six structural components that have been shown to be related to collective efficacy are operationalized by single indicators. For the second approach, we use exploratory factor analysis to reduce the dimensionality of the municipality-level data to four main dimensions that explain a large proportion of the observed variation: Residential instability, concentration of elderly people, concentration of previously partnered people, and socioeconomic status. Three of the dimensions ('residential instability', 'concentration of elderly people, and 'socioeconomic status') relate to some extent to the regional differentiators identified by Sampson and colleagues (1999) [13], but the 'concentration of previously partnered people' does not correspond to any of the theoretical components of regional differentiation.

With regard to our first research question concerning the theoretical components of regional differentiation, the findings suggest that none are consistently related to survey response outcomes in the ESS in Belgium. A less discouraging conclusion can be drawn for our second research question, regarding the main dimensions underlying the observed variation in the municipality data. One of the municipality dimensions relates to survey response outcomes in both rounds: Residential instability. Residential instability may hinder the development of social ties and a shared interest in the neighborhood [13], and people living in municipalities with a high residential turnover may therefore feel less inclined to help out when requested to participate in a survey. Residential instability may also hinder successfully reaching sample units because of locked gates barring entry to multi-unit buildings, people moving more frequently, and neighbors being less willing or less able, due to weak social ties, to give useful information on when and where to find them. When individuals' basic socio-demographics and interviewer-observed housing conditions and impediments to access are added, however, the effect of residential instability becomes less prominent.

Our overall conclusion is that the extent to which, and the way in which, aggregated data from official administrative registries helps to explain individuals' survey response outcomes is not only survey project specific, as suggested by Johnson and colleagues (2006) [12], but also round specific within one survey project. We find somewhat different results for round 7 of the ESS (fieldwork in 2014) compared with round 6 (fieldwork in 2012) in Belgium, even though the two rounds are iterations of the same overall survey project in that the key design features (mode, contact procedure, etc.) are identical and the topics covered very similar. The large amount of accessible municipality-level data from official administrative registries appears to have little relevance to explain survey response outcomes in round 6. By contrast, the round 7 results suggest that the available municipality-level data can be taken advantage of to explain individuals' survey response outcomes. What this appears to indicate is that there is no common set of regional differentiation indicators with a stable effect on nonresponse in the Belgian ESS. The observed differences in the regional correlates of survey nonresponse identified may be partially due to the differences in the fieldwork process. Round 6 went particularly well in Belgium, with the target net sample size being reached within the planned fieldwork period and with minimal refusal conversion efforts required. In round 7 in Belgium, in contrast, the target net sample size was not reached, even after extending the fieldwork period and making extensive refusal conversion efforts. The fieldwork process in round 7 being less smooth and homogeneous may explain why we find more, and different, regional correlates of survey nonresponse.

Regional differentiation explains only a part of the variability in survey response propensity. More generally, response in face-to-face surveys cannot be attributed exclusively to the potential respondents and their social environment. The interviewers also play an important role in actually getting people to participate (Groves & Couper, 1998). Both noncontact and refusal have been shown to be subject to interviewer effects (O'Muircheartaigh & Campanelli, 1999; Pickery & Loosveldt, 2002). As geographic proximity is an important factor in the assignment of sample units to interviewers in the Belgian ESS, the area effects in nonresponse cannot be straightforwardly separated from the possible interviewer effects.

The readily obtainable aggregated registry data appear to be of only limited usefulness in explaining and predicting survey nonresponse in the case of the ESS in Belgium. The few individual-level variables available from the sampling frame and the interviewer observations, on the other hand, show much more promise. The level of aggregation for which data is easily accessible from official administrative registries seems to be too high. The Belgian municipalities in our study are still relatively large aggregates and may be quite heterogeneous in both composition and context. Lower-level aggregated registry data may represent regional compositions and (local) social environments more accurately, but is more difficult to obtain. Individual-level registry data is even harder to procure but, if possible, may well be worth the effort. As the challenge of limited data availability from official registries at low levels is likely to persist, interviewer observations remain a valuable alternative.

Acknowledgements

This research was funded by FWO Flanders.

References

[1] de Leeuw E, de Heer W. Trends in household survey nonresponse: A longitudinal and international comparison. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, editors. Survey nonresponse. New York: John Wiley & Sons, Inc.; 2002. p. 41-54.

[2] Stoop I, Billiet J, Koch A, Fitzgerald R. Improving survey response: Lessons learned from the European Social Survey. Chichester, UK: John Wiley & Sons; 2010.

[3] Beullens K, Matsuo H, Loosveldt G, Vandenplas C. Quality report for the European Social Survey, Round 6. London: European Social Survey ERIC; 2014.

[4] Little RJA, Vartivarian S. Does weighting for nonresponse increase the variance of survey means? Surv Methodol. 2005;31(2):161-8.

[5] Sinibaldi J, Trappmann M, Kreuter F. Which Is the Better Investment for Nonresponse Adjustment: Purchasing Commercial Auxiliary Data or Collecting Interviewer Observations? Public Opin Q. 2014;78(2):440-73. doi 10.1093/poq/nfu003

[6] Groves RM, Couper M. Nonresponse in household interview surveys. New York: Wiley; 1998.

[7] Durrant GB, Steele F. Multilevel modelling of refusal and non-contact in household surveys: evidence from six UK government surveys. J Roy Stat Soc A Sta. 2009;172(2):361-81. doi 10.1111/j.1467-985X.2008.00565.x

[8] Groves RM, Cialdini RB, Couper MP. Understanding the decision to participate in a survey. Public Opin Q. 1992;56(4):475-95. doi 10.1086/269338

[9] House JS, Wolf SW. Effects of urban residence on interpersonal trust and helping behavior. J Pers Soc Psychol. 1978;36(9):1029-43.

[10] Couper MP, Singer E, Kulka RA. Participation in the 1990 decennial census: Politics, privacy, pressures. Am Polit Q. 1998;26(1):59-80.

[11] Abraham KG, Maitland A, Bianchi SM. Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does It Matter? Public Opin Q. 2006;70(5):676-703. doi 10.1093/poq/nfl037

[12] Johnson TP, Cho YI, Campbell RT, Holbrook AL. Using community-level correlates to evaluate nonresponse effects in a telephone survey. Public Opin Q. 2006;70(5):704-19. doi 10.1093/poq/nfl032

[13] Sampson RJ, Morenoff JD, Earls F. Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children. Am Sociol Rev. 1999;64:633-60.

[14] Matsuoka R, Maeda T. Neighborhood and individual factors associated with survey response behaviour: A multilevel multinomial regression analysis of a nationwide survey in Japan. SSJJ. 2015;18(2):217-32. doi 10.1093/ssjj/jyv011

[15] Van Goor H, Jansma F, Veenstra R. Differences in undercoverage and nonresponse between city neighbourhoods in a telephone survey. University of Groningen; 2005:867-78. (Psychological Reports). Report No. 96.

[16] European Social Survey. Round 6 Specification for Participating Countries. London: Centre for Comparative Social Surveys, City University London; 2011.

[17] European Social Survey. Round 7 Specification for Participating Countries. London: Centre for Comparative Social Surveys, City University London; 2013.

[18] European Social Survey Round 6 Data. Data file edition 2.1. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data for ESS ERIC; 2012.

[19] European Social Survey Round 7 Data. Data file edition 1.0. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data for ESS ERIC; 2014.

[20] Census 2011 België [Internet]. Statistics Belgium; 2015. Available from http://census2011.fgov.be/.

[21] Steele, F, Durrant, GB. Alternative approaches to multilevel modelling of survey non-contact and refusal: Approaches to modelling non-contact and refusal. Int Stat Rev. 2011;79(1):70-91. doi 10.1111/j.1751-5823.2011.00133.x

[22] Lynn P, Clarke P. Separating refusal bias and non-contact bias: evidence from UK national surveys. J Roy Stat Soc D-Sta. 2002; 51(3):319-33.

[23] Goldstein H, Rasbash J, Browne, WJ. Partitioning variation in multilevel models. Understanding Statistics. 2002;1:223-31. doi 10.1207/S15328031US0104_02

[24] O'Muircheartaigh C, Campanelli P. A multilevel exploration of the role of interviewers in survey non-response. J Roy Stat Soc A Sta. 1999; 162(3):437-446.

[25] Pickery J, Loosveldt GA. multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. Quality & Quantity. 2002;36:427-437.

Tables

Table 1: Final response outcome distribution and outcome rates, ESS round 6 and 7 in Belgium

	ESS6 Full sample N	ESS6 Sample in analysis N	ESS6 %	ESS7 Full sample N	ESS7 Sample in analysis N	ESS7 %
Respondent (completed interview)	1,869	1,869	57.72	1,769	1,769	55.37
Noncontact	209	190	5.87	172	166	5.20
Refusal	778	778	24.03	837	837	26.20
Not able and other nonresponse	326	323	9.98	324	323	10.11
Ineligible	85	78	2.41	102	100	3.13
	3,267	3,238		3,204	3,195	
Nonresponse rate			40.85			42.84
Noncontact rate			6.01			5.36
Conditional refusal rate			26.20			28.58

Table 2: Pearson correlation matrix of the six single indicators of regional differentiation

	(1)		(2)		(3)		(4)		(5)		(6)
(1) % changed address	1.00										
(2) % employers	-0.15	***	1.00								
(3) Low-income municipality	0.31	***	-0.11	**	1.00						
(4) Population density	0.30	***	-0.19	***	0.14	**	1.00				
(5) % non-Belgian nationality	0.56	***	-0.13	**	0.37	***	0.47	***	1.00		
(6) % younger than 15	0.22	***	-0.25	***	0.04		-0.19	***	0.12	**	1.00

* p < .05; ** p < .01; *** p < .001

Table 3: Municipality-level variables, descriptive statistics, and factor loadings

		Descriptives		Factor scores			
Name		**Mean**	**SD**	**F1**	**F2**	**F3**	**F4**
DENSITY	Natural logarithm of population density	5.73	1.16	0.61	0.16	-0.19	0.07
SMEDINKOM	Standardized median income	0.00	1.00	-0.35	-0.10	-0.12	0.72
IKAINKAS	Interquartile asymmetry in income	14.97	6.44	0.48	0.08	0.17	-0.10
DEMOPM	% men	0.49	0.01	-0.10	0.03	0.02	-0.05
PROPMIN15	% younger than 15	0.17	0.02	-0.08	-0.81	0.13	-0.04
PROP1524	% 15 to 24 years	0.12	0.01	0.07	-0.54	0.07	0.08
PROP4564	% 45 to 64 years	0.28	0.02	-0.35	0.47	-0.04	0.20
PROP65PL	% 65 years or older	0.17	0.03	0.03	0.93	0.15	0.04
GNSO	% secondary education not completed	0.15	0.04	-0.03	0.23	0.06	-0.84
SO	% secondary education	0.30	0.04	-0.59	0.38	-0.11	-0.19
HO	% higher education	0.25	0.07	-0.03	-0.13	-0.07	0.95
BELGNAT	% Belgian nationality	0.93	0.07	-0.66	0.30	0.15	0.16
EUMIGRAN	% EU if other nationality	0.70	0.16	-0.18	-0.25	0.15	-0.02
EUROPMIG	% other European if other nationality	0.07	0.06	-0.06	0.23	-0.08	-0.04
AFRIKMIG	% African if other nationality	0.10	0.08	0.32	-0.06	0.06	0.01
LATAMMIG	% Latin-American if other nationality	0.01	0.01	0.07	0.07	-0.20	0.26
NAMERMIG	% North-American if other nationality	0.01	0.04	-0.02	-0.10	0.07	0.16
AZIATMIG	% Asian if other nationality	0.10	0.08	0.10	0.41	-0.29	-0.06
HUWENWS	% married or legally cohabiting	0.44	0.04	-0.55	0.51	-0.63	0.01
SCHEID	% divorced or legally separated	0.08	0.01	0.52	0.13	0.64	-0.02
NOOITPAR	% never married or legally cohabiting	0.41	0.03	0.35	-0.83	0.31	0.09
WEDUWE	% widow(er)	0.06	0.01	0.04	0.47	0.75	-0.26
SINGLEHH	% single-person households	0.12	0.04	0.84	0.01	0.37	-0.11
ALJOKI	% single parents with young child(ren)	0.08	0.03	0.54	-0.40	0.61	-0.09
ALOUKI	% single parents with older child(ren)	0.02	0.00	0.01	-0.10	0.63	-0.16
PAARJOKI	% couples with young child(ren)	0.46	0.05	-0.67	-0.40	-0.39	0.19
PAAROUKI	% couples with older child(ren)	0.04	0.01	-0.45	0.25	-0.26	0.05
PAARGNKI	% couples without children	0.23	0.03	-0.28	0.85	-0.31	-0.02
UITWIJK1	% changed address	0.09	0.02	0.81	-0.27	0.27	-0.09
WERKZAAM	% employed	0.42	0.04	-0.61	0.19	-0.54	0.24
WERKLOOS	% unemployed	0.03	0.02	0.48	-0.40	0.58	-0.35
PENSIOEN	% retired	0.18	0.03	-0.05	0.92	0.32	0.00
STUDENT	% in education	0.08	0.01	-0.08	-0.42	0.06	0.67
HUISMV	% homemakers	0.12	0.03	0.72	-0.21	0.02	-0.34
WERKNMRS	% employees among employed	0.83	0.04	0.04	-0.15	0.14	-0.35
WERKGVRS	% employers among employed	0.05	0.01	-0.04	0.22	-0.10	0.40
WERKZLFS	% self-employed among employed	0.10	0.02	0.05	0.02	-0.14	0.43
ONBEWOON	% uninhabited buildings	0.12	0.09	0.34	0.29	0.16	-0.08
EENWOON	% in buildings with one unit	0.80	0.16	-0.86	-0.07	0.12	-0.08
TWEEWOON	% homes in buildings with two units	0.05	0.02	0.27	-0.44	0.16	0.11
DRIEWOON	% homes in buildings more units	0.14	0.15	0.86	0.15	-0.16	0.07
KOOPWOON	% in owned home	0.75	0.09	-0.89	0.16	-0.21	0.17
HUURWOON	% in rented home	0.22	0.09	0.92	-0.13	0.17	-0.17
EENKAMER	% homes with only one or two rooms	0.02	0.02	0.74	0.01	0.10	0.01
CTRWARM	% homes with central heating	0.76	0.10	-0.04	0.17	-0.54	0.49
BOUWWOI	% homes built before 1919	0.26	0.18	-0.05	-0.54	0.68	-0.06
BOUWWOIWOII	% homes built between 1919 and 1946	0.12	0.07	0.17	0.18	-0.04	-0.17
BOUWWOIINIEUW	% homes built between 1946 and 2006	0.58	0.16	0.02	0.51	-0.67	0.14
NIEUBOUW	% homes built after 2006	0.05	0.02	-0.31	0.10	-0.52	0.03

Table 4: Multilevel logistic regression coefficients of municipality variables of regional differentiation on survey nonresponse, noncontact, and refusal. ESS round 6

			Model 1				Model 2	
Nonresponse	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	4.6389	2.7019	0.0860		3.3319	2.8096	0.2357	
% employers	3.0509	3.3426	0.3614		3.9113	3.5237	0.2670	
Low-income municipality	-0.0886	0.1236	0.4737		-0.1300	0.1266	0.3041	
Population density	0.0165	0.0590	0.7803		-0.0293	0.0612	0.6321	
% non-Belgian nationality	0.2688	0.9956	0.7872		0.4239	1.0231	0.6786	
% younger than 15 years	-1.9936	2.5916	0.4417		-2.1432	2.7108	0.4292	
Individual level								
Female					0.0914	0.0753	0.2250	
Age 25 to 34 years					0.4735	0.1406	0.0008	***
Age 35 to 44 years					0.3130	0.1377	0.0231	*
Age 45 to 54 years					0.2516	0.1366	0.0655	
Age 55 to 64 years					0.3236	0.1411	0.0218	*
Age 65 or older					0.4911	0.1304	0.0002	***
Impediments to access					0.3439	0.0880	0.0001	***
Bad housing conditions					0.1875	0.0436	< 0.0001	***
			Model 1				**Model 2**	
Noncontact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	-2.4869	5.4294	0.6469		-5.7166	5.8640	0.3296	
% employers	5.0608	6.0823	0.4054		7.3867	6.6165	0.2642	
Low-income municipality	0.3891	0.2353	0.0982		0.3382	0.2449	0.1674	
Population density	0.1146	0.1228	0.3509		0.0637	0.1288	0.6206	
% non-Belgian nationality	1.3672	1.9312	0.4790		1.6807	2.0419	0.4105	
% younger than 15	2.2595	5.0563	0.6550		0.4448	5.3738	0.9340	
Individual level								
Female					-0.2814	0.1567	0.0726	
Age 25 to 34 years					0.4305	0.2520	0.0876	
Age 35 to 44 years					-0.0203	0.2670	0.9394	
Age 45 to 54 years					-0.0638	0.2677	0.8117	
Age 55 to 64 years					-0.0590	0.2783	0.8321	
Age 65 or older					-1.2860	0.3442	0.0002	***
Impediments to access					0.4295	0.1807	0.0175	*
Bad housing conditions					0.4061	0.0818	< 0.0001	***
			Model 1				**Model 2**	
Refusal, given contact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	3.0490	3.0941	0.3244		2.8380	3.1840	0.3727	
% employers	-0.2541	3.9184	0.9483		0.1265	4.0506	0.9751	
Low-income municipality	-0.3283	0.1455	0.0240	*	-0.3246	0.1468	0.0270	*
Population density	-0.0180	0.0681	0.7911		-0.0359	0.0697	0.6062	
% non-Belgian nationality	0.5737	1.1397	0.6147		0.5357	1.1582	0.6437	
% younger than 15 years	-1.5567	3.0418	0.6088		-1.2638	3.0822	0.6818	
Individual level								
Female					0.0923	0.0864	0.2856	
Age 25 to 34 years					0.4173	0.1684	0.0132	*
Age 35 to 44 years					0.4840	0.1623	0.0029	**
Age 45 to 54 years					0.4359	0.1609	0.0068	**
Age 55 to 64 years					0.3633	0.1672	0.0298	*
Age 65 or older					0.4409	0.1540	0.0042	**
Impediments to access					0.1631	0.1005	0.1046	
Bad housing conditions					-0.0257	0.0511	0.6147	

Table 5: Multilevel logistic regression coefficients of municipality variables of regional differentiation on survey nonresponse, noncontact, and refusal. ESS round 7

			Model 1				Model 2	
Nonresponse	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	4.7301	2.7921	0.0902		3.2849	2.7916	0.2393	
% employers	5.6524	3.5993	0.1163		4.5249	3.6319	0.2128	
Low-income municipality	-0.1505	0.1392	0.2794		-0.1439	0.1374	0.2951	
Population density	0.2116	0.0587	0.0003	***	0.1653	0.0588	0.0049	**
% non-Belgian nationality	-1.0938	0.9147	0.2318		-0.9312	0.9090	0.3056	
% younger than 15 years	-6.7366	2.8586	0.0184	*	-6.0337	2.8689	0.0355	*
Individual level								
Female					0.1522	0.0767	0.0472	*
Age 25 to 34 years					0.7008	0.1427	< 0.0001	***
Age 35 to 44 years					0.6173	0.1409	< 0.0001	***
Age 45 to 54 years					0.4219	0.1394	0.0025	**
Age 55 to 64 years					0.3835	0.1434	0.0075	**
Age 65 or older					0.6615	0.1342	< 0.0001	***
Impediments to access					0.4340	0.0863	< 0.0001	***
Bad housing conditions					0.0528	0.0472	0.2632	
			Model 1				**Model 2**	
Noncontact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	9.2584	5.4136	0.0872		5.1611	5.7458	0.3691	
% employers	4.0728	6.3359	0.5203		4.2007	6.7462	0.5335	
Low-income municipality	-0.2643	0.2517	0.2938		-0.3237	0.2611	0.2151	
Population density	0.2046	0.1150	0.0754		0.1201	0.1196	0.3152	
% non-Belgian nationality	-0.6159	1.7712	0.7280		-0.0989	1.8435	0.9572	
% younger than 15 years	-3.7146	4.8393	0.4427		-5.0285	5.1002	0.3242	
Individual level								
Female					0.0586	0.1663	0.7246	
Age 25 to 34 years					0.6124	0.2495	0.0141	*
Age 35 to 44 years					-0.1552	0.2810	0.5808	
Age 45 to 54 years					-0.3138	0.2900	0.2791	
Age 55 to 64 years					-0.9745	0.3574	0.0064	**
Age 65 or older					-1.0648	0.3323	0.0014	**
Impediments to access					0.7552	0.1820	< 0.0001	***
Bad housing conditions					0.4148	0.0919	< 0.0001	***
			Model 1				**Model 2**	
Refusal, given contact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
% changed address	4.7987	2.8182	0.0886		4.4896	2.8159	0.1109	
% employers	4.6830	3.5078	0.1819		4.1561	3.5030	0.2354	
Low-income municipality	-0.0672	0.1377	0.6253		-0.0518	0.1370	0.7051	
Population density	0.1309	0.0591	0.0268	*	0.1084	0.0599	0.0702	
% non-Belgian nationality	-2.3304	0.9554	0.0147	*	-2.4169	0.9589	0.0117	*
% younger than 15 years	-9.3636	2.8781	0.0011	**	-8.4356	2.8662	0.0032	**
Individual level								
Female					0.1242	0.0846	0.1420	
Age 25 to 34 years					0.5249	0.1633	0.0013	**
Age 35 to 44 years					0.5642	0.1583	0.0004	***
Age 45 to 54 years					0.4020	0.1574	0.0106	*
Age 55 to 64 years					0.4916	0.1594	0.0020	**
Age 65 or older					0.2669	0.1530	0.0812	
Impediments to access					0.2419	0.0938	0.0099	
Bad housing conditions					-0.1492	0.0533	0.0051	**

Table 6: Comparison of survey response outcome models with municipality variables selected based on theory

	AIC			BIC			LR-test (P-value)	
Outcome	**Model 0**	**Model 1**	**Model 2**	**Model 0**	**Model 1**	**Model 2**	**Model 1 versus Model 0**	**Model 2 versus Model 1**
ESS round 6								
Nonresponse	4257	4259	4221	4269	4307	4318	0.1349	< 0.0001
Noncontact	1427	1427	1370	1439	1475	1467	0.0595	< 0.0001
Refusal, given contact	3399	3404	3404	3411	3452	3500	0.3749	0.0431
ESS round 7								
Nonresponse	4170	4149	4098	4183	4197	4195	< 0.0001	< 0.0001
Noncontact	1280	1274	1203	1292	1322	1299	0.0063	< 0.0001
Refusal, given contact	3489	3471	3452	3501	3519	3547	< 0.0001	< 0.0001

Note: Model 0 does not include any explanatory variables, Model 1 includes the municipality-level single indicators of regional differentiation, Model 2 adds the individual-level age, gender, and interviewer observations of impediments to access and the physical condition of housing

Table 7: Multilevel logistic regression coefficients of municipality dimensions on survey nonresponse, noncontact, and refusal. ESS round 6

			Model 1				Model 2	
Nonresponse	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.0883	0.0361	0.0143	*	0.0247	0.0388	0.5248	
Concentration of elderly people	0.0077	0.0465	0.8688		0.0010	0.0476	0.9824	
Concentration of previously partnered people	0.0437	0.0483	0.3655		0.0585	0.0497	0.2389	
Socioeconomic status	0.0527	0.0495	0.2863		0.0764	0.0506	0.1306	
Individual level								
Female					0.0889	0.0753	0.2380	
Age 25 to 34 years					0.4786	0.1406	0.0007	***
Age 35 to 44 years					0.3141	0.1378	0.0226	*
Age 45 to 54 years					0.2532	0.1366	0.0637	
Age 55 to 64 years					0.3247	0.1411	0.0214	*
Age 65 or older					0.4913	0.1304	0.0002	***
Impediments to access					0.3468	0.0884	0.0001	***
Bad housing conditions					0.1837	0.0436	< 0.0001	***

			Model 1				Model 2	
Noncontact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.1114	0.0664	0.0936		0.0083	0.0733	0.9103	
Concentration of elderly people	-0.2213	0.0994	0.0261	*	-0.1996	0.1040	0.0549	
Concentration of previously partnered people	0.1509	0.0924	0.1025		0.1685	0.0973	0.0833	
Socioeconomic status	0.0379	0.0922	0.6813		0.0998	0.0961	0.2987	
Individual level								
Female					-0.2831	0.1564	0.0702	
Age 25 to 34 years					0.4215	0.2515	0.0937	
Age 35 to 44 years					-0.0181	0.2665	0.9457	
Age 45 to 54 years					-0.0705	0.2675	0.7921	
Age 55 to 64 years					-0.0575	0.2779	0.8360	
Age 65 or older					-1.2737	0.3440	0.0002	***
Impediments to access					0.4935	0.1792	0.0059	**
Bad housing conditions					0.4021	0.0815	< 0.0001	***

			Model 1				Model 2	
Refusal, given contact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.0158	0.0426	0.7111		-0.0073	0.0450	0.8712	
Concentration of elderly people	-0.0148	0.0544	0.7850		-0.0236	0.0548	0.6671	
Concentration of previously partnered people	0.0326	0.0565	0.5634		0.0483	0.0573	0.3988	
Socioeconomic status	0.0639	0.0581	0.2709		0.0645	0.0586	0.2705	
Individual level								
Female					0.0911	0.0864	0.2916	
Age 25 to 34 years					0.4278	0.1684	0.0111	*
Age 35 to 44 years					0.4862	0.1623	0.0027	**
Age 45 to 54 years					0.4384	0.1610	0.0065	**
Age 55 to 64 years					0.3642	0.1672	0.0294	*
Age 65 or older					0.4409	0.1540	0.0042	**
Impediments to access					0.1713	0.1012	0.0906	
Bad housing conditions					-0.0321	0.0511	0.5303	

Table 8: Multilevel logistic regression coefficients of municipality dimensions on survey nonresponse, noncontact, and refusal. ESS round 7

			Model 1				Model 2	
Nonresponse	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.1976	0.0365	< 0.0001	***	0.1434	0.0374	0.0001	***
Concentration of elderly people	0.1562	0.0483	0.0012	**	0.1358	0.0482	0.0048	**
Concentration of previously partnered people	-0.0535	0.0494	0.2783		-0.0356	0.0493	0.4701	
Socioeconomic status	0.1813	0.0487	0.0002	***	0.1786	0.0485	0.0002	***
Individual level								
Female					0.1572	0.0765	0.0400	*
Age 25 to 34 years					0.7046	0.1426	< 0.0001	***
Age 35 to 44 years					0.6203	0.1409	< 0.0001	***
Age 45 to 54 years					0.4243	0.1393	0.0023	**
Age 55 to 64 years					0.3819	0.1433	0.0077	**
Age 65 or older					0.6580	0.1342	< 0.0001	***
Impediments to access					0.4232	0.0860	< 0.0001	***
Bad housing conditions					0.0566	0.0469	0.2281	
			Model 1				**Model 2**	
Noncontact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.2310	0.0614	0.0002	***	0.1142	0.0672	0.0892	
Concentration of elderly people	-0.0119	0.0888	0.8931		0.0036	0.0932	0.9692	
Concentration of previously partnered people	-0.0242	0.0988	0.8067		-0.0206	0.1032	0.8421	
Socioeconomic status	0.1935	0.0892	0.0300	*	0.2303	0.0920	0.0123	*
Individual level								
Female					0.0616	0.1663	0.7111	
Age 25 to 34 years					0.6217	0.2496	0.0128	*
Age 35 to 44 years					-0.1494	0.2810	0.5948	
Age 45 to 54 years					-0.3116	0.2901	0.2827	
Age 55 to 64 years					-0.9764	0.3577	0.0063	**
Age 65 or older					-1.0607	0.3323	0.0014	**
Impediments to access					0.7559	0.1834	< 0.0001	***
Bad housing conditions					0.4191	0.0918	< 0.0001	***
			Model 1				**Model 2**	
Refusal, given contact	**Estimate**	**SE**	**P-value**		**Estimate**	**SE**	**P-value**	
Municipality level								
Residential instability	0.0954	0.0360	0.0080	**	0.0709	0.0374	0.0584	
Concentration of elderly people	0.2276	0.0475	< 0.0001	***	0.2166	0.0477	< 0.0001	***
Concentration of previously partnered people	-0.0415	0.0503	0.4088		-0.0157	0.0508	0.7566	
Socioeconomic status	0.1040	0.0493	0.0347	*	0.0970	0.0494	0.0495	*
Individual level								
Female					0.1284	0.0844	0.1283	
Age 25 to 34 years					0.5305	0.1632	0.0012	**
Age 35 to 44 years					0.5700	0.1582	0.0003	***
Age 45 to 54 years					0.4061	0.1573	0.0098	**
Age 55 to 64 years					0.4922	0.1593	0.0020	**
Age 65 or older					0.2658	0.1530	0.0823	
Impediments to access					0.2331	0.0936	0.0127	
Bad housing conditions					-0.1490	0.0532	0.0051	**

Table 9: Comparison of survey response outcome models with municipality dimensions based on exploratory factor analysis

	AIC			BIC			LR-test (P-value)	
Outcome	**Model 0**	**Model 1**	**Model 2**	**Model 0**	**Model 1**	**Model 2**	**Model 1 versus Model 0**	**Model 2 versus Model 1**
ESS round 6								
Nonresponse	4257	4256	**4218**	**4269**	4292	4303	0.0671	< 0.0001
Noncontact	1427	1423	**1366**	**1439**	1459	1450	0.0175	< 0.0001
Refusal, given contact	**3399**	3405	3404	**3411**	3441	3488	0.7712	0.0369
ESS round 7								
Nonresponse	4170	4134	**4083**	4183	4170	**4168**	< 0.0001	< 0.0001
Noncontact	1280	1269	**1197**	1292	1305	**1281**	0.0008	< 0.0001
Refusal, given contact	3489	3467	**3447**	**3501**	3503	3531	< 0.0001	< 0.0001

Note: Model 0 does not include any explanatory variables, Model 1 includes the municipality dimensions, Model 2 adds the individual-level age, gender, and interviewer observations of impediments to access and the physical condition of housing.